JOURNAL OF PHYSICAL ORGANIC CHEMISTRY, VOL. 11, 508–511 (1998)

Review Commentary

A perspective on biological catalysis

Stephen J. Benkovic

Pennsylvania State University, Department of Chemistry, 414 Wartik Laboratory, University Park, Pennsylvania 16802, USA

Received 18 September 1997; accepted 5 January 1998

Most discussions of biological catalysis commence with the concept of transition state stabilization that postulates that the increased rate for enzymic reactions is inversely proportional to the decrease in the free energy of the transition state. In its unadorned version, the concept does not consider the possible contribution to catalysis from ground-state destabilization—the difference in free energy between the ground and transition states is all that matters—nor does it describe how and when along the reaction coordinate the interaction between catalyst and substrate arises given the femtosecond lifetimes of transition states.

that were additive or non-additive, meaning that substitution of a pair of amino acids showed them either to act independently ($\Delta G_i = 0$) or to interact ($\Delta G_i \neq 0$) with one another. For the examples of L28F–L54F and L28Y–L54F, the first pair exhibited $\Delta G_i = 0$ for key steps in the cycle, whereas the second pair showed $\Delta G_i \neq 0$ particularly for the steps involving cofactor binding and hydride transfer in the forward but not the reverse direction (Fig. 2). The simplest interpretation of this effect on the chemical step in one direction but not the other is a specific interaction within the enzyme–substrate–cofactor ternary complex increasing its ground

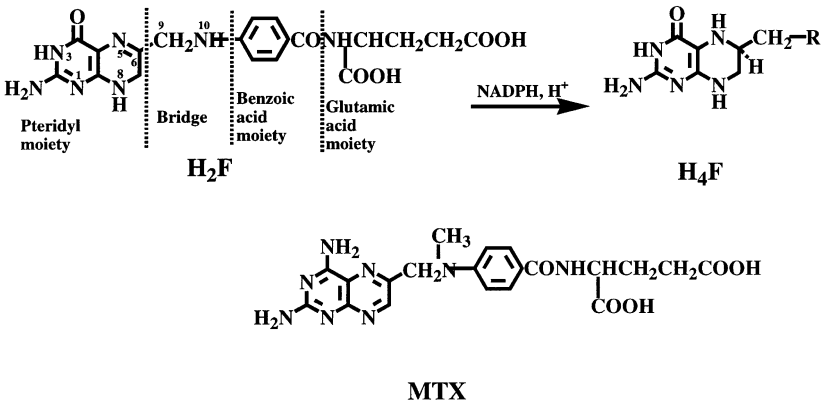

Figure 1. The reaction catalyzed by dihydrofolate reductase

The enzyme dihydrofolate reductase (Fig. 1) has been used as a paradigm to probe these issues given the wealth of structural and kinetic information available.[1,2] The method used was site-specific mutagenesis, particularly of hydrophic residues that line the active site. Sets of single and double mutant proteins were constructed and their kinetic properties for key steps in the turnover cycle were determined.[3] Of particular interest is whether the mutations led to changes in the free energy for such steps

**Correspondence to:* Stephen J. Benkovic, Pennsylvania State University, Department of Chemistry, 414 Wartik Laboratory, University Park, Pennsylvania 16802, USA.
E-mail: sjb1@psu.edu

state reactivity that is absent at the transition state and the product ternary complex. The fact that ΔG_i had different values for the various steps suggests that the enzyme undergoes a variety of conformational changes throughout the reaction cycle.

Physical evidence for the conformational flexibility within various regions of the enzyme was gained by multi-dimensional NMR relaxation measurements of the protein backbone. Previous assignments[4] facilitated measurement of the longitudinal and transverse relaxation times of the N—H amide bond for each of the approximately 160 amino acids in the binary folate complex. These in turn were transformed into τ_e, R_{ex} and S^2 terms, which are related to backbone motions on the ps

CCC 0894–3230/98/080508–04 $17.50

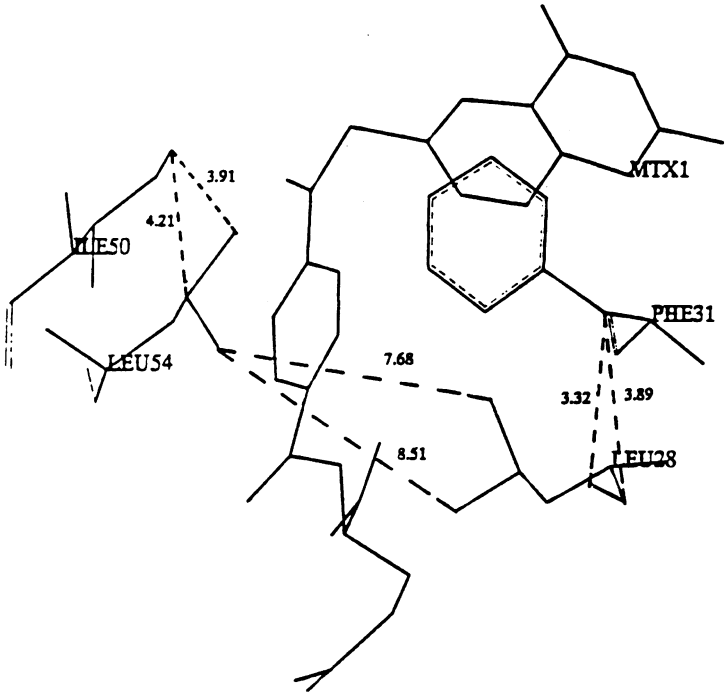

Figure 2. Distances (Å) between the side-chain of Leu28, Phe31, Ile50 and Leu54 in the binary complex of *E. coli* dihydrofolate reductase with MTX

and μs time-scales, and also the amplitude of the motion.[5] Of particular interest were residues within the helices and β-sheet that surround the active site which exhibited motions on the μs time-scale, and also amino acids within loop 1 and the βF–βG loop whose dynamics were on the ps time-scale and likewise showed the highest amplitudes (Fig. 3). One interpretation views the lower frequency motions as indicative of conformational changes involving elements of secondary structure that optimize the loci of the cofactor and substrate within the reaction volume of the active site for maximum reactivity. This placement is further refined by the closure of the two loops whose conformational flexibility is reflected by high frequency and amplitude motions relative to the peptide backbone. In this manner the population of substrate and cofactor conformations within the active site is selected for high reactivity. Mutations within loop 1 produced less active enzymes with changed rates of hydride transfer by factors up to 500-fold.[6] Mutations within the outer βF–βG loop (either deletions of G121 or its replacement by valine) led to enzymes which not only showed rates of hydride transfer that were 200–400-fold slower but also exhibited conformational changes which were now largely rate limiting (1–20 s–1) on the pathways leading up to the final ternary substrate–cofactor–enzyme complex. The effects on the latter loop are particularly surprising since the residues within this region are >20 Å from the folate binding site and do not physically contact the substrate in any of the relevant x-ray crystallographic structures. One can conclude that active sites have a multitude of near and remote side-chain interactions, some modulated by substrate linking residues within the active site, a molecular recognition redundancy within the active site to preserve function and a conformational flexibility in active site elements to achieve optimum chemical catalysis and turnover.

How, then, does nature construct these exquisite catalysts? An examination of the active site of the enzyme glycinamide ribonucleotide transformylase focused on three residues, N106, N108 and D144, which had been implicated as key to its catalytic function by active site labeling verified by x-ray crystal structures. These three residues were mutated into every possible combination of the 20 amino acids through random, saturation mutagenesis.[7] The constructs containing the mutated genes were used to transform auxotrophic *Escherichia coli* cells and the survivors were isolated and sequenced. In the 120 clones found in a library of *ca* 88 000 transformants, no genes were found that contained two or three replacements for the wild-type sequence. Only the wild-type gene and single changes were found able to complement the auxotrophic cell, inferring that this constellation of residues has been optimized within this structure for transformylase activity (Fig. 4)

The importance of these three residues and their relative spacing to the enzyme's capability for transferring a formyl residue from N^{10}-formyl tetrahydrofolate to an acceptor substrate is underscored by searching the genomic base for this sequence. This ensemble appears in genes encoding enzymes that use the same cofactor for

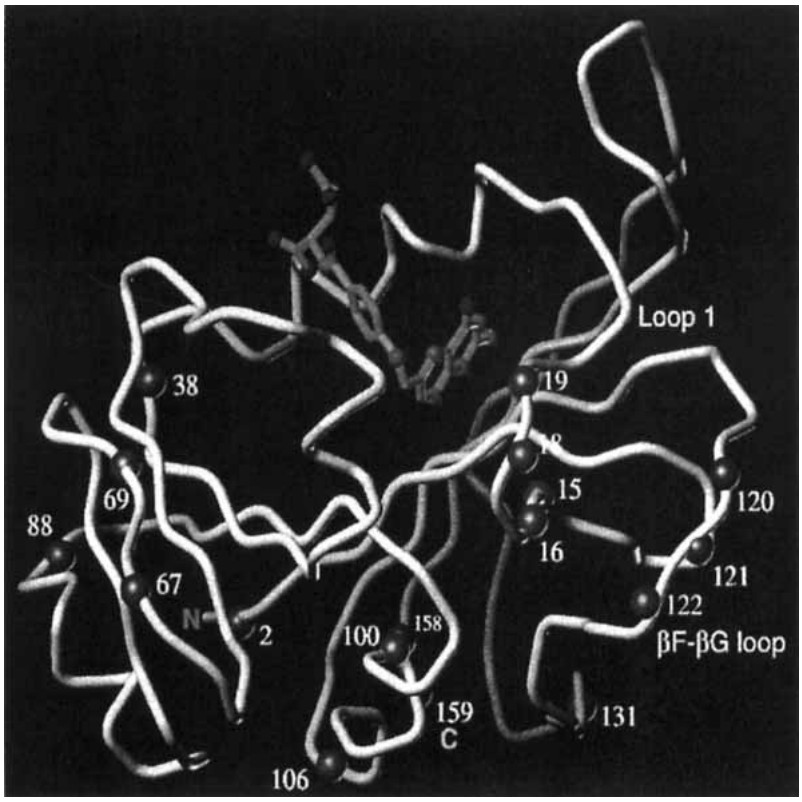

Figure 3. Map of the effective correlation time (τ_c) for internal motions, on the picosecond to nanosecond time-scale, for the backbone amide nitrogens of the DHFR–folate complex. Residue-specific τ_c and generalized S^2 values (in parentheses) are I12, 6051 ± 2081 ps (0.65); G15,81 ± 24 ps (0.88); M16, 3894 ± 5438 ps (0.63); N18, 1128 ± 296 ps (0.73); A19, 1595 ± 325 ps (0.65); K38, 1720 ± 210 ps (0.59); G67, 1610 ± 120 ps (0.29); D69, 729 ± 136 ps (0.58); V88, 2350 ± 692 ps (0.52); Y100, 4904 ± 3188 ps (0.81); K106, 1061 ± 266 ps (0.63); E120, 100 ± 425 ps (0.72); G121, 614 ± 48 ps (0.50); D122, 30 ± 6 ps (0.68); D131, 40 ± 30 ps (0.87); R158, 32 ± 13 ps (0.85); and R159, 1111 ± 320 ps (0.74)

other formyl transfer reactions. It would appear that this sequence is an essential element of a super family of enzymes whose common mechanistic feature is to catalyze a formyl transfer reaction from this cofactor. One can speculate that all these reactions proceed through a putative tetrahedral species whose formation and decay require a proton transfer and that these three residues can optimize the free energy requirements of this process.

The attempted construction of catalytic entities from biomaterials serves as a valuable exercise to test our understanding of the underlying principles and possibly, to reformulate them. In experiments influenced by the need for conformational flexibility and by the concept of

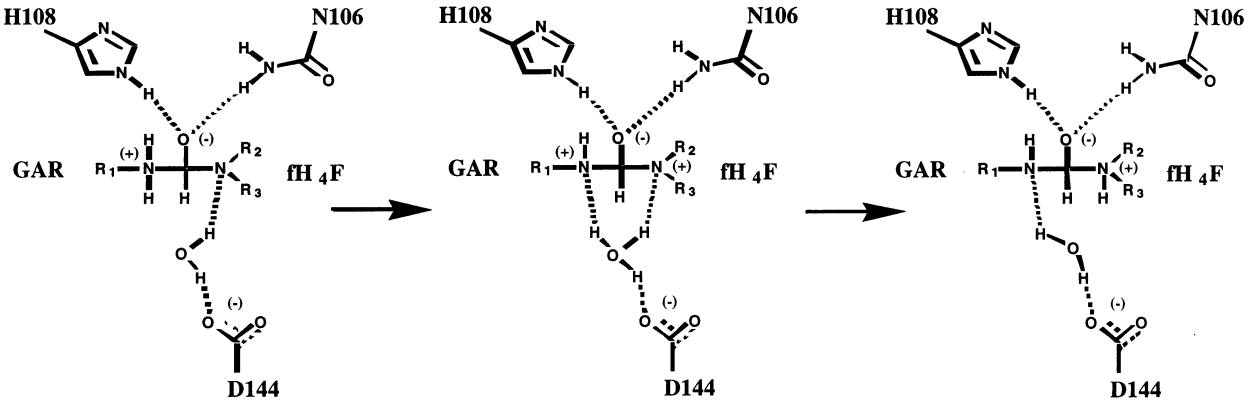

Figure 4. The proton shuttle around the putative tetrahedral intermediates in the reaction catalyzed by glycinamide ribonucleotide transformylase

modular design, gene segments of existing binding or catalytic sites have been joined combinatorially in an effort to create chimeric enzymes.

The enzymes encoded by the *E. coli* genes *pur*U and *pur*N, N^{10}-formyl tetrahydrofolate hydrolyase and glycinamide ribonucleotide (GAR) transformylase, respectively, catalyze similar yet distinct reactions.[8] N^{10}-Formyl tetrahydrofolate hydrolase uses water to cleave N^{10}-formyl tetrahydrofolate into tetrahydrofolate and formate, whereas GAR transformylase catalyses the transfer of formyl from N^{10}-formyl tetrahydrofolate to GAR to yield formyl-GAR and tetrahydrofolate. The two enzymes show significant homology (*ca* 60%) in the carboxyl-terminal region which, from the GAR transformylase crystal structure and labeling studies, is known to be the site of N^{10}-formyl tetrahydrofolate binding. Both enzymes possess three appropriately spaced residues at their active sites implicated in catalyzing the transfer of the formyl group. Hybrid proteins were created by joining varying length segments of the *N*-terminal region of the *pur*N gene (GAR binding region) and the *C*-terminal (N^{10}-formyl tetrahydrofolate binding) region of *pur*U. Active *pur*N/*pur*U hybrids were then selected for their ability to complement an auxotrophic *E. coli* strain. Hybrids able to complement the auxotrophs were purified to homogeneity and assayed for activity. The specific activity of two hybrid proteins was within 100–1000-fold of the native *pur*N GAR transformylase, validating the approach of constructing an enzyme active site from functional parts of others.

In conclusion, these studies among other things underscore the importance of conformational flexibility in enzymic catalysis. The linking of the dynamics of specific regions in the protein to catalytic events is a goal of future work, and its relative importance will greatly impact the developing field of *de novo* enzyme design.

REFERENCES

1. C. Bystroff and J. Kraut. *Biochemistry* **30,** 2227 (1991).
2. C. A. Fierke, K. A. Johnson and S. J. Benkovic. *Biochemistry* **26,** 4085–4092 (1987).
3. C. R. Wagner, Z. Huang, S. F. Singleton and S. J. Benkovic. *Biochemistry* **34,** 15671–15680 (1995).
4. C. J. Falzone, J. Cavanagh, M. Cowart, A. G. Palmer, C. R. Matthews, S. J. Benkovic and P. E. Wright. *J. Biolmol. NMR* **4,** 349–366 (1994).
5. D. M. Epstein, S. J. Benkovic and P. E. Wright. *Biochemistry* **34,** 11037–11048 (1995).
6. L. Li, C. J. Falzone, P. E. Wright and S. J. Benkovic. *Biochemistry* **31,** 7826–7833 (1992).
7. M. S. Warren and S. J. Benkovic. *Protein Eng.* **10,** 63–68 (1997).
8. A. E. Nixon, M. S. Warren and S. J. Benkovic. *Proc. Natl. Acad. Sci. USA* **94,** 1069–1073 (1997).

